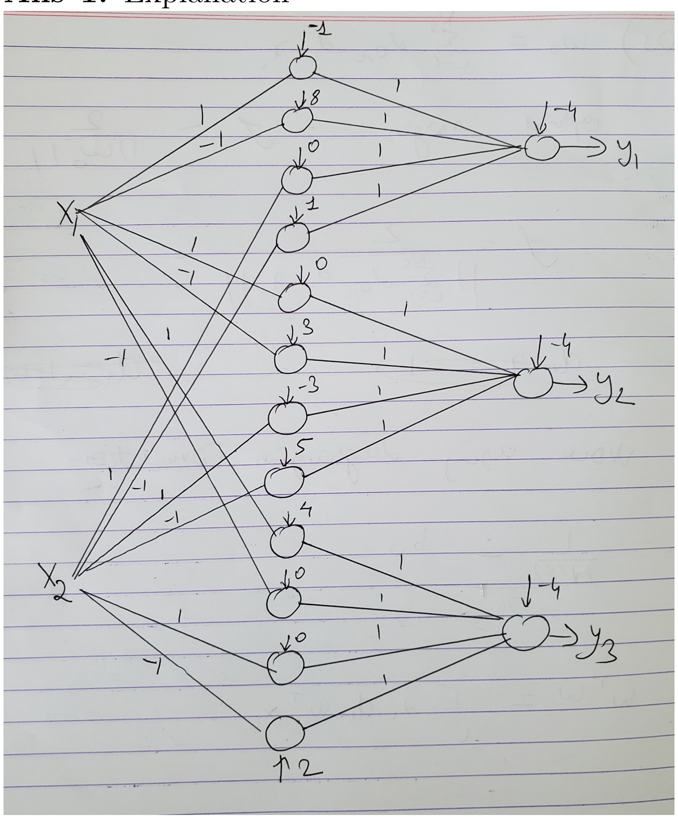
Homework Assignment 4

Rushabh Ashok Dharia
CSCI-B555 Machine Learning

April 27, 2019

Note: Homework was discussed with Varun Miranda

Ans 1. Explanation



As mentioned in the question we need to design a Neural Network with 3 output units we can consider that Class 4 is recognized when y_i = -1 for $1 \leq i \leq 3$
This means that when all neurons in the last layer do not fire (or output -1) we can say that the point belongs to class 4. This can be easily showed by adding another layer which will take the outputs y_1, y_2 and y_3 as inputs

Consider the first 4 neurons in the 1st Layer

1st Neuron will output 1 if $x_1 \geq 1$, otherwise 0
2nd Neuron will output 1 if $x_1 \leq 8$, otherwise 0
3rd Neuron will output 1 if $x_2 \geq 0$, otherwise 0
4th Neuron will output 1 if $x_2 \leq 1$, otherwise 0
Now, the 1st Neuron in the 2nd Layer will output 1 if all 4 neurons mentioned above output 1, else it will output 0
Similarly, the other 2 classes can be determined

Ans 2. *a).*

$$
\begin{aligned}
J(u_j) &= \sum_{j=1}^{K}\sum_{i=1}^{N} w_{ij}||x_i - u_j||^2 \\
&= \sum_{j=1}^{K}\sum_{i=1}^{N} w_{ij}(x_i^2 + u_j^2 - 2x_iu_j) \\
&= \sum_{j=1}^{K}\sum_{i=1}^{N} w_{ij}x_i^2 + w_{ij}u_j^2 - 2w_{ij}x_iu_j \\
&\text{Taking derivative w.r.t } u_j \text{ and setting it equal to 0} \\
\frac{dJ(u_j)}{du_j} &= \sum_{j=1}^{K}\sum_{i=1}^{N} 0 + 2w_{ij}u_j - 2w_{ij}x_i = 0 \\
\sum_{j=1}^{K}\sum_{i=1}^{N} w_{ij}uj &= \sum_{j=1}^{K}\sum_{i=1}^{N} w_{ij}x_i \\
\text{Therefore, } \hat{u_j} &= \frac{\sum_{i=1}^{N} w_{ij}x_i}{\sum_{i=1}^{N} w_{ij}}
\end{aligned}
$$

□

b). Numerator denotes the points in each cluster and the denominator denotes the number of points in each cluster. The fraction as a whole denotes the weighted average of all points in the cluster

□

Ans 3. *a).*

$$
\begin{aligned}
y(i) &= \sum_{j=1}^{K} w_j(n) e^{\frac{||x_i - u_j(n)||^2}{2\sigma^2(n)}} \\
\frac{\partial y(i)}{\partial w_j(n)} &= \sum_{j=1}^{K} e^{\frac{||x_i - u_j(n)||^2}{2\sigma^2(n)}} \\
\frac{\partial y(i)}{\partial u_j(n)} &= w_j(n) e^{\frac{||x_i - u_j(n)||^2}{2\sigma^2(n)}} . \frac{x_i - u_j(n)}{\sigma^2(n)} \\
\frac{\partial y(i)}{\partial \sigma(n)} &= w_j(n) e^{\frac{||x_i - u_j(n)||^2}{2\sigma^2(n)}} . \frac{(x_i - u_j(n))^2}{\sigma^3(n)} \\
\text{Now, } E &= \frac{1}{2} \sum_{i=1}^{N} (d(i) - y(i))^2 \\
\frac{\partial E}{\partial w_j(n)} &= - \sum_{i=1}^{N} (d(i) - y(i)) . \frac{\partial y(i)}{\partial w_j(n)} \\
&= - \sum_{i=1}^{N} (d(i) - y(i)) \sum_{j=1}^{K} e^{\frac{||x_i - u_j(n)||^2}{2\sigma^2(n)}} \\
\frac{\partial E}{\partial u_j(n)} &= - \sum_{i=1}^{N} (d(i) - y(i)) . \frac{\partial y(i)}{\partial u_j(n)} \\
&= - \sum_{i=1}^{N} (d(i) - y(i)) \sum_{j=1}^{K} w_j(n) e^{\frac{||x_i - u_j(n)||^2}{2\sigma^2(n)}} . \frac{x_i - u_j(n)}{\sigma^2(n)} \\
\frac{\partial E}{\partial \sigma(n)} &= - \sum_{i=1}^{N} (d(i) - y(i)) . \frac{\partial y(i)}{\partial \sigma(n)} \\
&= - \sum_{i=1}^{N} (d(i) - y(i)) \sum_{j=1}^{K} w_j(n) e^{\frac{||x_i - u_j(n)||^2}{2\sigma^2(n)}} . \frac{||x_i - u_j(n)||^2}{\sigma^3(n)}
\end{aligned}
$$

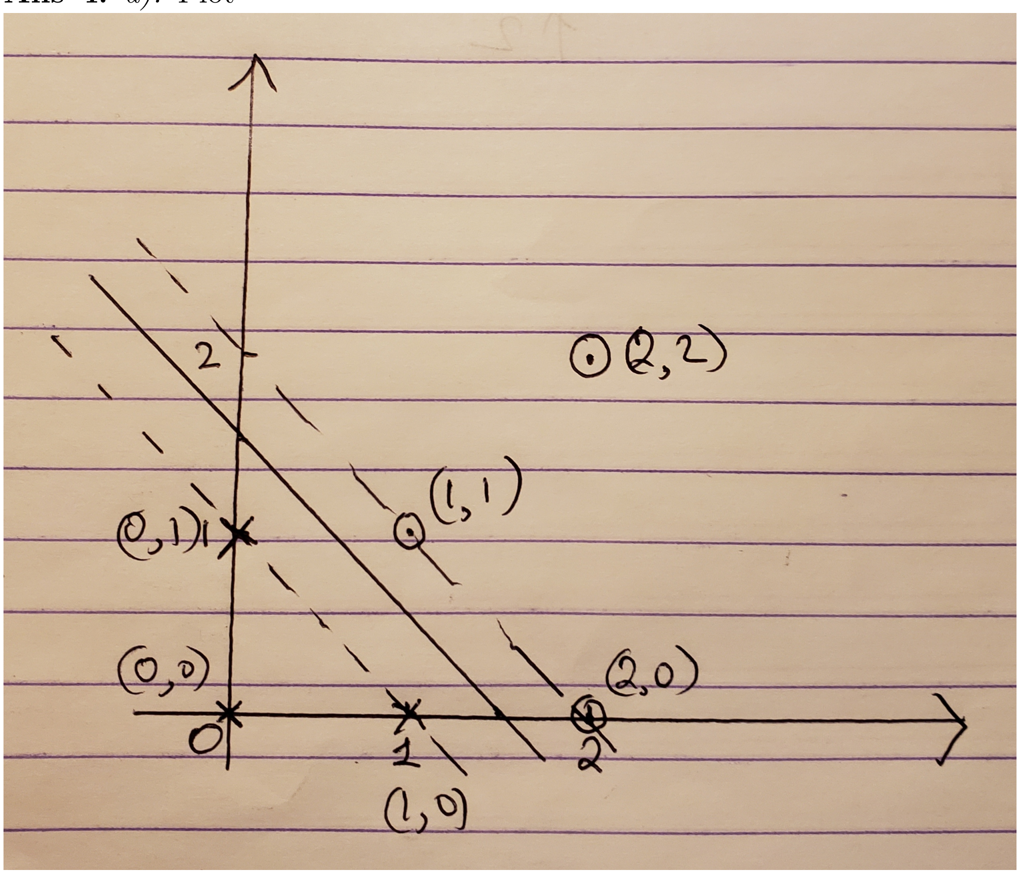
$\square$

b).

$$
\begin{aligned}
w_{kj}(n+1) &= w_{kj}(n) - \eta_w \frac{\partial E}{\partial w_{kj}} \\
&= w_{kj}(n) + \eta_w \sum_{i=1}^{N} (d(i) - y(i)) \sum_{j=1}^{K} e^{\frac{||x_i - u_j(n)||^2}{2\sigma^2(n)}} \\
w_{kj}(n+1) &= w_{kj}(n) - \eta_u \frac{\partial E}{\partial u} \\
&= w_{kj}(n) + \eta_u \sum_{i=1}^{N} (d(i) - y(i)) \sum_{j=1}^{K} w_j(n) e^{\frac{||x_i - u_j(n)||^2}{2\sigma^2(n)}} . \frac{x_i - u_j(n)}{\sigma^2(n)} \\
w_{kj}(n+1) &= w_{kj}(n) - \eta_\sigma \frac{\partial E}{\partial \sigma} \\
&= w_{kj}(n) - \eta_\sigma \sum_{i=1}^{N} (d(i) - y(i)) \sum_{j=1}^{K} w_j(n) e^{\frac{||x_i - u_j(n)||^2}{2\sigma^2(n)}} . \frac{||x_i - u_j(n)||^2}{\sigma^3(n)}
\end{aligned}
$$

$\square$

Ans 4. *a).* Plot



$$
\begin{aligned}
[w_1 \ \ w_2]^T[2 \ \ 0] + b_0 &= 1 \\
[w_1 \ \ w_2]^T[1 \ \ 1] + b_0 &= 1 \\
2w_1 + b_0 &= 1 \qquad (1) \\
w_1 + w_2 + b_0 &= 1 \qquad (2) \\
\text{Subtracting (2) from (1)} \\
w_1 - w_2 &= 0 \\
w_1 &= w_2 \qquad (3)
\end{aligned}
$$

$$
\begin{aligned}
[w_1 \ \ w_2]^T[1 \ \ 0] + b_0 &= -1 \\
w_1 + b_0 &= -1 \qquad (4) \\
\text{Subtracting (1) from (4)} \\
-w_1 &= -2 \\
w_1 &= 2 \qquad (5) \\
w_2 &= 2 \qquad \text{Using (3) and (5)} \\
b_0 = -3 \qquad \text{Using (4) and (5)}
\end{aligned}
$$

□

b).

The Support Vectors are (0,1), (1,0), (1,1) and (2,0

□

Ans 5.

$$J(w,b,\alpha) = \frac{w^T w}{2} - \sum_i^N \alpha_i[d_i(w^T x + b) - 1]$$

$$\phi(w) = \frac{1}{2} w^T w$$

$$\text{Now, } \phi(w) = J(w,b,\alpha)$$

$$\text{Therefore, } \sum_i^N \alpha_i[d_i(w^T x + b_0) - 1] = 0$$

$$\sum_i^N \alpha_i d_i w^T x + \sum_i^N \alpha_i d_i b_0 - \sum_i^N \alpha_i = 0 \qquad (1)$$

$$\text{Now, } w = \sum_i^N \alpha_i d_i x$$

$$\text{Therefore, } w^T w = \sum_i^N \alpha_i d_i w^T x \qquad (2)$$

$$\sum_i^N \alpha_i d_i = 0 \qquad (3)$$

Substituting (2) and (3) in (1)

$$w^T w + 0 - \sum_i^N \alpha_i = 0$$

$$w^T w = \sum_i^N \alpha_i$$

$$||w|| = (\sum_i^N \alpha_i)^{1/2} \qquad (4)$$

$$\text{Now, } \rho = \frac{2}{||w||}$$

$$\text{Therefore, } \rho = \frac{2}{(\sum_i^N \alpha_i)^{1/2}} \qquad \text{Using, (4)}$$